Nickos Papadatos

Description of Research

(a) PhD Thesis

[0] Contribution to the theory of order statistics and to approximation of distributions.

The thesis is divided in two parts. First part (chapters 1–3) considers properties of ordered samples, while in the second part (chapters 4–6), some bounds and approximation results for the total variation distance among probability measures are derived.

Specifically, in the first chapter we define the *intermediate order statistics* and it is shown, [1], that they can be used as estimators of the population quantiles. In second chapter we derive some variance bounds for order statistics using a method of Cacoullos and Papathanasiou, *Statist. Probab. Lett.* (1985)**3**, 175–184 – [5]. In the third chapter we obtain the best possible upper bound for the variance of a single order statistic, in terms of the population variance, [2].

The second part of the thesis is devoted to find convenient upper bounds for the total variation distance between two probability measures. Specifically, in fourth chapter we obtain some information-type bounds between an arbitrary distribution and an extreme value distribution. Fifth chapter extends these results to the general case, and the typical bound is of the form

$$d_{TV}(X,Y) \le c_Y \, \mathbb{E} \left| \frac{f'(X)}{f(X)} - \frac{g'(X)}{g(X)} \right|,$$

where $d_{TV}(X,Y) = \sup_A |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|$ is the total variation distance, f, g are the densities of X, Y, and c_Y is a constant depending only on Y – see [4]. Finally, the sixth chapter contains results using different functionals, the so called w-functions of X and Y. It turns out that the last mentioned bounds have better behavior than their information-type counterparts of chapter five, [3].

(b) Research Papers

(i) In International Journals / Refereed Volumes

[1] Intermediate order statistics with applications to nonparametric estimation.
The uniform *intermediate* order statistics are constructed by a method that uses the original observations of the sample and an exterior generation of suitable Beta r.v.'s. Moreover, a method to approximate the intermediate order statistics in the general (non-uniform) case is given. Numerical simulations show that these r.v's have well behavior as estimators of the quantiles, even for small values of the sample size.

[2] Maximum variance of order statistics.
If $\sigma^2 > 0$ is the population variance, it is proved that the variance of the standardized order statistic, $X_{i:n}/\sigma$, cannot exceed a constant $c_{i:n}$. The proof of the result is based on Hoeffding's identity, and the bound is best possible (attainable).

[3] Distance in variation between two arbitrary distributions via the associated w-functions.
The main result in Cacoullos, Papathanasiou and Utev, *Ann. Probab.* (1994)**22**, 1607–1618 is extended, obtaining bounds for the total variation distance of arbitrary r.v.'s X and Y. The main result has the form (provided that the first two moments are identical)

$$d_{TV}(X,Y) \le c_Y \mathbb{E}\left|1 - \frac{w_X(X)}{w_Y(X)}\right|,$$

where the function $w_X(\cdot)$ is defined in the (assumed) interval support of X by

$$w_X(x) = \frac{1}{\sigma^2 f(x)} \int_{-\infty}^{x} (\mu - t) f(t) dt, \quad \mu = \mathbb{E}X, \quad \sigma^2 = \mathrm{Var}X,$$

and similarly for w_Y. These functions appear in variance bounds and in Stein-type identities, so they are called covariance kernels.

It should be noted that Cacoullos, Papathanasiou and Utev proved this result only for a normal r.v. Y. The above general bound is extended to the discrete case, and also to the case where X is a sum of possible dependent r.v.'s. Applications to specific examples provide the rate of convergence to the limiting distribution.

[4] Distance in variation and a Fisher-type information.
We derive an information-type bound for the total variation distance and the result is also extended to the discrete case. A simple application of the main inequality simplifies a result of Barron, *Ann. Probab.* (1986)**14**, 336–342, regarding a strengthened CLT (L^1 convergence of densities), in the sense of entropy. Other applications are related to the extreme value theory, providing explicit results to the rate of convergence.

[5] A generalization of variance bounds.
Upper and lower variance bounds for order statistics, involving the *density–quantile* function $f(F^{-1}(\cdot))$, are obtained. A simple application simplifies the known result for the limiting variance.

[6] A note on maximum variance of order statistics from symmetric populations.
When the population is symmetric, the bounds in [2] are no longer tight. Under symmetry, we obtain the new sharp bounds, verifying great improvement compared to the general case.

[7] Exact bounds for the expectations of order statistics from non-negative populations.
We derive the best possible upper bound for the expectation of a single order statistic when the population is non-negative and has mean $\mu > 0$. Also, we derive the corresponding result for the difference of two order statistics (generalized spacings). In practice, e.g., in Reliability Systems, the units are lifetimes, hence non-negative. Thus, the assumption of non-negativity is not too restrictive. Since the bounds do not depend on the population variance, we gain a major improvement to the classical results

of Hartley–David–Gumbel, *Ann. Math. Statist.* (1954)**25**, 85–99, 75–84, and Moriguti, *Ann. Math. Statist.* (1953)**24**, 107–113.

[8] Variance inequalities for covariance kernels and applications to central limit theorems.
We prove a convolution-type inequality for the standardized sum of two independent absolutely continuous r.v.'s. Based on this inequality (which is valid under the restriction that the support of the r.v.'s is a finite or infinite interval), it is shown that the rate of convergence (in total variation) of the standardized sums $(S_n - n\mu)/(\sigma\sqrt{n})$ (from i.i.d. r.v.'s $X, X_1, X_2, \ldots$) to the standard normal, Z, is at least $O(n^{-1/2})$. Moreover, an explicit value of the constant c_X for which

$$d_{TV}((S_n - n\mu)/(\sigma\sqrt{n}), Z) \le c_X/\sqrt{n}$$

is given. The result applies also to the multivariate case. It should be noted that the proof is extremely simple.

[9] Total variation distance and generalized covariance kernels.
We provide an extended version of [3], in order to obtain total variation bounds in terms of a function $Z_f(\cdot; h(\cdot))$, where f is the density of X and $h(\cdot)$ is an arbitrary function. The function $Z_f(\cdot; h(\cdot))$ is a generalized covariance kernel, since the identity function $h(x) = x$ leads to $Z_f(\cdot; x) = w_X(\cdot)$, as in [3]. When h is strictly increasing, the bound takes the form

$$d_{TV}(X, Y) \le c_Y \mathbb{E}\left|1 - \frac{Z_f(X; h)}{Z_g(X; h)}\right|,$$

where the constant c_Y depends only on Y. Using $h = -g'/g$, where g is the density of Y, it is shown that, under general conditions,

$$d_{TV}(X_n, Y) \to 0 \text{ if and only if } Z_{f_n}(X_n; -g'/g) \to 1 \text{ in probability.}$$

[10] Variational inequalities for arbitrary multivariate distributions.
We present bounds for the total variation in the multivariate setup and in the discrete case. The bounds are similar to those given in [3], and extend the main result from Papathanasiou, *J. Multivariate Anal.* (1996)**58**, 189–196 to the general (not necessarily normal) case. Interesting applications include discrete distributions where the limiting distribution has independent components. In such cases, the bound simplifies considerably and we can investigate the rate of convergence. As illustrative examples, we obtain rates of convergence of the Multinomial and the Negative Multinomial to the multivariate Poisson.

[11] Three elementary proofs of the Central Limit Theorem with applications to random sums.
The present work was presented in the Conference in the Memory of Stamatis Cambanis,

18–19 December 1995 at Athens, Greece, and contains three elementary proofs and CLT. Moreover, it includes a result regarding the case of random sums, i.e., sums of N i.i.d. r.v.'s where N is an r.v. with values in $\mathbb{N}$.

[12] Upper bound for the covariance of extreme order statistics from a sample of size three.
Let $X_{1:3} \le X_{2:3} \le X_{3:3}$ be the order statistics based on a sample of size $n = 3$ from an arbitrary distribution function with mean μ and variance $\sigma^2 > 0$. Using the orthogonal polynomial system of Legendre on $[0,1]$, it shown that

$$\operatorname{Cov}[X_{1:3}, X_{3:3}] \le \frac{6}{a^2}\sigma^2,$$

where $a \simeq 0.16838$ is the unique positive root to the equation $\tanh(a/2) = a/6$, and the equality characterizes the *hyperbolic sine distribution* with density

$$f(x) = \frac{1/a}{\sqrt{(x-\mu)^2 + \lambda^2\sigma^2}}, \qquad |x-\mu| < a\sigma\sqrt{\frac{2}{a^2-24}},$$

where $\lambda = \sqrt{2(36-a^2)/(a^2-24)} \simeq 0.25089$.

[We point out that the same inequality can be derived if we apply some results from the theory of integral operators].

[13] Expectation bounds on linear estimators from dependent samples.
Let $(X_1, X_2, \ldots, X_n)'$ be a vector of arbitrary r.v's (possibly dependent with possibly different marginals), and consider their order statistics $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$. Assuming $\mu_i = \mathbb{E}X_i$ and $\sigma_i^2 = \operatorname{Var} X_i < \infty$, $i = 1, 2, \ldots, n$, it is shown that for any constants $c_1, c_2, \ldots, c_n$,

$$\sum_{i=1}^{n} c_i(\mathbb{E}X_{i:n} - \overline{\mu}) \le \left(\sum_{i=1}^{n}(a_i - \overline{c})^2\right)^{1/2} \left(\sum_{i=1}^{n}\left\{(\mu_i - \overline{\mu})^2 + \sigma_i^2\right\} - n\operatorname{Var}\overline{X}\right)^{1/2},$$

where $\overline{\mu} = n^{-1}\sum_{i=1}^{n}\mu_i$, $\overline{c} = n^{-1}\sum_{i=1}^{n} c_i$, $\overline{X} = n^{-1}\sum_{i=1}^{n} X_i$ and $(a_1, a_2, \ldots, a_n)'$ is the ℓ^2-projection of $(c_1, c_2, \ldots, c_n)'$ onto the convex cone of component-wise increasing vectors of $\mathbb{R}^n$ (in particular, $a_i = c_i$ for all i if and only if the c_i's are non-decreasing in i). A similar lower bound holds.

The bound is best possible whenever the X_i's are exchangeable. Moreover, it provides an essential improvement over the bounds given by Arnold and Groeneveld, *Ann. Statist.* (1979)**7**, 220–223, Aven, *J. Appl. Probab.* (1985)**22**, 723–728 and Lefèvre, *Stochastic Anal. Appl.* (1986)**4**, 351-356, as it becomes clear from its applications.

[14] Distribution and expectation bounds on order statistics from possibly dependent variates.
We define a new dependence structure that is useful especially for reliability systems. Specifically, the n-variate random vector $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)'$ is called Maximally Stable of order j, MAS(j) for short, if the distribution $F_{(j)}$ of $\max\{X_{k_1}, \ldots, X_{k_j}\}$ is invariant for any subset $\{k_1, \ldots, k_j\} \subseteq \{1, \ldots, n\}$ of order j. For example, all n-variate random vectors are MAS(n), while the random vectors that are MAS(1) are exactly those with the same marginals, $F_{(1)}$ say. Note that exchangeable vectors are MAS(j) for all $j \in \{1, \ldots, n\}$.

Denote by $X_{k:n}$ and $F_{k:n}$ the k-th order statistic and its distribution. It is shown that if $\boldsymbol{X}$ is MAS(j) then for $k \geq j$,

$$F_{k:n}(x) \leq \min\left\{1, \frac{(n)_j}{(k)_j} F_{(j)}(x)\right\} \text{ and } \mathbb{E}X_{k:n} \geq \frac{1}{a}\int_0^a F_{(j)}^{-1}(u)\mathrm{d}u,$$

where $a = (k)_j/(n)_j$, $(c)_j = c(c-1)\cdots(c-j+1)$, and $F_{(j)}^{-1}(u) = \inf\{x : F_{(j)}(x) \geq u\}$, $0 < u < 1$, is the left-continuous inverse of $F_{(j)}$. Equality is always attainable for all j, k, n, and for any given distribution function $F_{(j)}$. Similar results hold for vectors that are *Minimally Stable of order* j, MIS(j), i.e., when the distribution of $\min\{X_{k_1}, \ldots, X_{k_j}\}$ is the same for all j-tuples. We point out the above bounds for MAS(j) and MIS(j) systems are applicable to reliability systems without restricting the components of the system to be independent or identically distributed; the only requirement is that a MAS(j) or a MIS(j) condition is fulfilled. This happens because $1 - F_{k:n}$ and $\mathbb{E}X_{k:n}$ is the reliability and the expected time to failure of an $(n+1-k)$-out-of-n system, respectively, and the condition MAS(j) reads as "parallel homogeneity of order j of the components".

[15] Unified variance bounds and a Stein-type identity.
This work was presented in the Conference in honor of Professor Theophilos Cacoullos, 3–6 June 1999 at Athens, Greece. It is shown that for any absolutely continuous r.v. X with finite variance σ^2, there exists a unique r.v. X^* (viewed as a transformation on X), such that the *generalized Stein identity* is fulfilled, that is,

$$\mathrm{Cov}[X, g(X)] = \sigma^2 \mathbb{E} g'(X^*)$$

for any absolutely continuous function g defined on the convex hull of the support of X, provided that $\mathbb{E}|g'(X^*)| < \infty$. The properties of this transformation are investigated in detail, and it is shown that the same r.v. X^* appears to both upper and lower bounds for $\mathrm{Var}\, g(X)$. Moreover, we investigate the inverse transform, and we show a convolution-type identity, which in the case where $X_1, X_2, \ldots, X_n$ are i.i.d., simplifies to

$$(X_1 + X_2 + \cdots + X_n)^* \stackrel{\mathrm{d}}{=} X_1^* + X_2 + \cdots + X_n.$$

[16] An application of a density transform and the local limit theorem.
We present a generalization of [3], showing that for any two r.v.'s X and Y, with densities f, g, means μ, m, and variances σ^2, s^2,

$$d_{TV}(X,Y) \le 2\int \left| f(x) - \frac{\sigma^2 g(x)}{s^2 g^*(x)} f^*(x) \right| dx + c_Y|\mu - m|,$$

where f^*, g^*, are the densities of X^*, Y^* (see [15]), and the constant c_Y can be chosen as $c_Y = 2/\mathbb{E}|Y-m|$. In the interesting case where $Y = Z$ is standard normal and $\sigma = s$, the bound simplifies to

$$d_{TV}((X-\mu)/\sigma, Z) \le 3d_{TV}(X, X^*) + \frac{\sqrt{\pi}}{\sigma\sqrt{2}}|\mu - m|,$$

and it is shown that for a sequence of absolutely continuous r.v.'s X_n with means $\mu_n \to \mu$ and variances $\sigma_n^2 \to \sigma^2 > 0$,

$$d_{TV}((X_n-\mu)/\sigma, Z) \to 0 \text{ if and only if } d_{TV}(X_n, X_n^*) \to 0.$$

Based on this result and using the convolution-type identity of [15], a simple proof of the local limit Theorem in its full generality, i.e., without assuming an interval support, is given. The original result is due to Prohorov, *Dokl. Acad. Nauk. SSSR* (1952)**83**, 797-800 (in Russian). We note here that at least three proofs of this Theorem appeared during the last decades (Barron, *Ann. Probab.* (1986)**14**, 336–342, Mayer–Wolf, *Ann. Probab.* (1990)**18**, 840–850, Cacoullos, Papathanasiou and Utev, *Ann. Probab.* (1994)**22**, 1607–1618); however, none of them treats the completely general case.

[17] The use of spacings in the estimation of a scale parameter.
The Best Linear Unbiased Estimators, BLUE, where defined by Lloyd, *Biometrika* (1952)**39**, 89–95. They are linear estimators of the form $\sum_{i=1}^n c_i X_{i:n}$, where $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ is the order sample, unbiased for the location or the scale parameter of a location-scale family, and they have minimal variance. In the present paper we use the spacings $Z_i = X_{i+1:n} - X_{i:n}$, $i = 1, 2, \ldots, n-1$, in order to obtain a simple alternative form for these estimators. The simplified version enable us to provide a partial positive answer to the open question if the scale estimator is non-negative with probability one. Note that the question remains unsolved in its full generality. Moreover, our technique provides simple forms for the estimators when we deal with censored samples from the Uniform location-scale family.

[18] Poisson approximation for a sum of dependent indicators: an alternative approach.
We introduce the notion of Totally Negatively Dependent (TND for short) r.v.'s as follows: the r.v.'s $X_1, X_2, \ldots, X_n$ are called TND if for each $i = 1, 2, \ldots, n$, the r.v.'s X_i and $X^{(i)} = \sum_{j \ne i} X_j$ are Negatively Quadrant Dependent, i.e., if $\mathrm{Cov}[f(X_i), g(X^{(i)})] \le 0$ for any pair of non-decreasing functions $f, g : \mathbb{R} \to \mathbb{R}$, for which the covariance is finite.

Assume that $X_1, X_2, \ldots, X_n$ are 0–1 indicators, $\mathbb{E}X_i = p_i = \mathbb{P}[X_i = 1]$, and set $W = \sum_{i=1}^n X_i$, $\mu = \mathbb{E}W$ and $\sigma^2 = \mathrm{Var}W$. One of the results is the following: If $X_1, X_2, \ldots, X_n$ are TND and $\mathcal{P}_\lambda$ denotes a Poisson r.v. with mean $\lambda \geq \mu$,

$$d_{TV}(W, \mathcal{P}_\lambda) \leq (1 - e^{-\lambda}) \left(1 - \frac{\sigma^2}{\lambda}\right) + \min\left\{1, \frac{(2/e)^{1/2}}{\lambda^{1/2}}\right\} (\lambda - \mu),$$

where $d_{TV}(X, Y)$ denotes the total variation distance of X and Y.

The proof is based on a refinement of the methodology used in [16], and the bound extends a classic result in Poisson approximation (see Barbour, Holst and Janson, *Poisson Approximation* (Oxford Studies Prob. **2**), Oxford University Press, 1992, Corollary 2.C.2) to TND 0–1 indicators. Note that the classic result was shown only for Negatively Related (NR) 0–1 indicators, and that the class of NR 0–1 indicators is strictly smaller than the class of TND 0–1 indicators.

The paper includes an application to a generalized birthday problem, investigating the rate of convergence. Moreover, some results comparing different notions on negative relations are given.

[19] Bounds on expectation of order statistics from a finite population.
Consider a simple random sample $X_1, X_2, \ldots, X_n$, arising from sampling without replacement from a finite ordered population $\Pi = \{x_1 \leq x_2 \leq \cdots \leq x_N\}$, and let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be the corresponding ordered sample ($n \leq N$). Define $\mu = N^{-1} \sum_{i=1}^N x_i$ and $\sigma^2 = N^{-1} \sum_{i=1}^N (x_i - \mu)^2$ to be the population mean and variance, respectively, noting that the X_i's have mean μ and variance σ^2, but they are not independent (they are, merely, exchangeable).

In the present article we obtain the best possible upper and lower bounds for $\mathbb{E}X_{i:n}$ and $\mathbb{E}[X_{n:n} - X_{1:n}]$ in terms of μ and σ^2, and we characterize the populations that attain the equality in the bounds. Similar results are obtained for the covariance in the simplest case where $n = 2$. An interesting future is that, as $N \to \infty$, the bounds (and the corresponding optimal populations) approximate the classic well-known results for the i.i.d. case.

[20] Multivariate covariance identities with an application to order statistics.
We show multivariate covariance identities of the form

$$\mathrm{Cov}[h^j(\boldsymbol{X}), g(\boldsymbol{X})] = \mathbb{E}[z^j(\boldsymbol{X}) g_j(\boldsymbol{X})], \ j = 1, 2, \ldots, n,$$

where $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)'$ is an absolutely continuous random vector, $g : \mathbb{R}^n \to \mathbb{R}$ a function with partial derivatives $g_j(\boldsymbol{x}) = \partial g(\boldsymbol{x})/\partial x_j$, and each $z^j : \mathbb{R}^n \to \mathbb{R}$ is a function depending on the density of $\boldsymbol{X}$ and the given function $h^j : \mathbb{R}^n \to \mathbb{R}$.

These identities generalize the results of Cacoullos and Papathanasiou, *J. Multivariate Anal.* (1992)**43**, 173–184, and improve the so called, Siegel's identity, Siegel, *J. Amer. Statist. Assoc.* (1993)**88**, 77–80. Some applications to order statisics arizing from an arbitrary multivariate normal are also given.

[21] Bounds on expectations of L-statistics from without replacement samples. Consider an ordered sample $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ from a finite ordered population of size $N \ge n$, taken without replacement – see [19]. Let

$$L = L(c_1, c_2, \ldots, c_n) = \sum_{i=1}^{n} c_i X_{i:n}$$

be the general form of an L-Statistic. Applying a suitable projection technique, we obtain upper and lower bounds for $\mathbb{E}L$. These bounds are, in most cases, optimal. The results are applied to some interesting cases, including trimmed means.

[22] Heteroscedastic one-way ANOVA and lack of fit tests.
It is well-known that in the one-way ANOVA case, the statistic $F = \mathrm{MST}/\mathrm{MSE}$ has asymptotically, under the null hypothesis $H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$, a χ^2 distribution, provided that the number of observations in every cell tends to infinity, while the number of cells, say a, remains bounded. In the present work we assume that the number of cells, a, tends to infinity, while the number of observations in each cell remains bounded. This situation is common in practical problems with many levels, and also applies to lack-of-fit tests, where, e.g., an unknown link function should be estimated from a small number of observations falling in a small neighborhood of its domain, and this situation is very common even if the sample size is large.

In the present work we apply a projection method in order to simplify the investigation of the asymptotic distribution of the statistic $\sqrt{a}\,(\mathrm{MSE})(F-1) = \sqrt{a}\,(\mathrm{MST} - \mathrm{MSE})$, as $a \to \infty$. Under general conditions it is shown that the statistic converges to a particular normal distribution. We also investigate the heteroscedastic case (the variances are different in every cell), as well as the unbalanced case where the number of observations is allowed to be different from one cell to another. Moreover, we obtain the asymptotic distribution under local alternatives. The power of the test as well as the rate of convergence to its limiting distribution is studied via simulations.

[23] The q-factorial moments of discrete q-distributions and a characterization of the Euler distribution.
For $0 < q < 1$ and $x \in \mathbb{R}$, the q-number of x is defined by $[x]_q = (1 - q^x)/(1 - q)$, Similarly, the q-factorial of order k of x is defined as $[x]_{k,q} = [x]_q [x-1]_q \cdots [x-k+1]_q$. These numbers are related to the so called q-distributions, appearing in the study of sums of n independent 0–1 indicators having different success probabilities (q-binomial, which as, $n \to \infty$, converges to Heine distribution). They also appear in the study of the number of failures before the k-th success (q-Pascal, which converges to Euler distribution as $k \to \infty$). In the present work we define the k-th order q-factorial moment of a random variable X taking values in $\mathbb{N}$, and we give expressions that connect the q-factorial moments with the usual factorial moments of X. The expressions are applied to q-distributions, showing that their q-factorial moments are simple, in contrast to their (usual) moments. Moreover, a moment characterization of the Euler distribution is given. According to this, if $\mathbb{E}[X_\lambda]_{2,q} = \{\mathbb{E}[X_\lambda]_q\}^2$ for every λ in a power-series family X_λ then X follows the Euler distribution.

[24] Characterizations of discrete distributions using the Rao-Rubin condition. This work has been presented in the conference 5th Lattice Path Combinatorics and Discrete Distributions, June 2002, Athens, Greece, and contains characterizations of discrete distributions. Specifically, assume that $(N_1, N_2, \ldots, N_k)$ is a random vector with values in $\mathbb{N}^k$, and that the Rao–Rubin partial independence condition is satisfied, namely,

$$\mathbb{P}[N_2 = n_2 | N_1 = 0] = \mathbb{P}[N_2 = n_2], \; n_2 \in \mathbb{N}.$$

Using a well-known Lemma of Shanbhag, *J. Appl. Probab.* (1977)**14**, 640–646, it is shown that if

$$\mathbb{P}[N_1 = n_1, N_2 = n_2 \ldots, N_k = n_k | N = n] = \frac{1}{c(n)} \prod_{j=1}^{k} a_j(n_j),$$

where $N = N_1 + \cdots + N_k$, and $c, a_1, \ldots, a_k : \mathbb{N} \to \mathbb{R}^+$ are arbitrary, then the r.v.'s $N_1, N_2, \ldots, N_k$ are independent and they follow distributions of particular forms.

Applying the result we obtain some characterizations of Poisson and Negative Binomial, similar to those given by Rao and Srivastava, *Sankhyā Ser. A* (1979)**41**, 124–128.

[25] On Rychlik's expectation bound for L-estimates based on identically distributed variates.
This work provides a simple convenient proof of the well-known result of Rychlik, *Statistics* (1993)**24**, 9–15, regarding optimal bounds on expectations of L-statistics, $L = \sum_{i=1}^{n} c_i X_{i:n}$, from possibly dependent identically distributed samples $X_1, X_2, \ldots, X_n$, in terms of the common marginal distribution F of X_i. The key idea goes as follows: Consider an r.v. $I(j,n)$, uniformly distributed in $\{j, \ldots, n\}$ and independent of the X_i's. Furthermore, consider another r.v. $U(j,n)$, uniformly distributed in the interval $[(j-1)/n, 1]$. Setting $F^{-1}(u) = \inf\{x : F(x) \geq u\}$, $0 < u < 1$, it is proved that the r.v.'s

$$X = X_{I(j,n):n} \;\text{ and }\; Y = F^{-1}(U(j,n))$$

are stochastically ordered: $X \leq_{st} Y$. Therefore, it follows at once that

$$\frac{1}{n-j+1} \sum_{i=j}^{n} \mathbb{E} X_{i:n} = \mathbb{E} X \leq \mathbb{E} Y = \frac{n}{n-j+1} \int_{(j-1)/n}^{1} F^{-1}(u) du.$$

The above inequality yields Rychlik's result. The present technique generalizes to other situations; e.g., when we cannot assume identical marginals for $X_1, X_2, \ldots, X_n$.

[26] The discrete Mohr and Noll inequality with applications to variance bounds. Mohr and Noll, *Math. Nachr.* (1952)**7**, 55–59, obtained an interestind extention of the Cauchy-Schwarz inequality, $(\int_a^b g(t)dt)^2 \leq (b-a) \int_a^b (g'(t))^2 dt$, involving higher order derivatives of g. In the present paper we obtain a discrete analogue of this inequality, replacing derivatives with forward differences. Using the discrete inequality, we obtain variance bounds for an arbitrary function $g(X)$ of an integer-valued r.v. X. The

complicated bounds are simplified considerably when X belongs to the Ord family of distributions (discrete Pearson system). In this case, the typical bound has the form

$$(-1)^n \operatorname{Var} g(X) \le (-1)^n \sum_{k=0}^{n} \frac{(-1)^k}{(k+1)! \prod_{j=0}^{k}(1-j\delta)} \mathbb{E} q^{[k+1]}(X)(\Delta^{k+1} g(X))^2,$$

where $q^{[k+1]}(x) = q(x)q(x+1)\dots q(x+k)$, $\Delta^{k+1}g$ is the $(k+1)$-th forward difference of g, and $q(x) = \delta x^2 + \beta x + \gamma$ is the characteristic polynomial of the probability mass function p of X, for which

$$\sum_{j \le k} (\mu - j)p(j) = p(k)q(k), \quad k \in \mathbf{Z} \quad (\mu = \mathbb{E}X).$$

We note that the above bounds are applicable in statistical problems. As an illustrative example, we obtain upper and lower bounds for the variance of the UMVU estimator of $\log p$ in samples from the Geometric distribution with parameter p, showing its asymptotic efficiency.

It is worth pointing out that the discrete inequality, proved in the present paper, is strictly better than the original Mohr and Noll inequality – see Remark 3.1 and, for more details, see the corresponding Technical Report.

[27] An extended Stein-type covariance identity for the Pearson family, with applications to lower variance bounds.
We investigate a class of Bessel-type variance bounds for functions of discrete and continuous Pearson r.v.'s. These bounds are based on the orthogonal polynomials P_k obtained from Rodrigues formula,

$$P_k(x) = \frac{(-1)^k}{p(x)} \Delta^k[q^{[k]}(x-k)p(x-k)], \quad \text{or} \quad P_k(x) = \frac{(-1)^k}{f(x)} \frac{d^k}{dx^k}[q(x)^k f(x)],$$

for discrete (with probability mass function p) or continuous (with density f) Pearson r.v. First, it is shown that the following inversion formula holds:

$$\begin{aligned} q(x)^k f(x) &= \frac{(-1)^k}{(k-1)!} \int_{-\infty}^{x} (x-y)^{k-1} P_k(y) f(y) dy \\ &= \frac{1}{(k-1)!} \int_{x}^{+\infty} (y-x)^{k-1} P_k(y) f(y) dy; \end{aligned}$$

a similar result is true for the discrete case. Using these inversion formulae, we show that the Fourier coefficients of g can be expressed in terms of its derivatives/differences, namely

$$\mathbb{E} P_k(X) g(X) = \mathbb{E} q(X)^k g^{(k)}(X), \quad \text{or} \quad \mathbb{E} P_k(X) g(X) = \mathbb{E} q^{[k]}(X) \Delta^k g(X),$$

provided that X has finite $2k$-th moment and the RHS's are finite. The lower bound (which is, just, a finite form of Bessel's inequality) takes the form

$$\operatorname{Var} g(X) \ge \sum_{k=1}^{n} \frac{\mathbb{E}^2 q(X)^k g^{(k)}(X)}{k! \, \mathbb{E} q(X)^k \prod_{j=k-1}^{2k-2}(1-j\delta)},$$

and it is quite similar to the corresponding bound in [26], obtained from a Mohr and Noll-type inequality. We emphasize that in applying the bound *it is not necessary to have an infinite sequence of orthogonal polynomials* (it is not even necessary for X to have moments of any order); in fact, the inequality is valid for any fixed n, provided that $2n$ moments exist. Hence, the bound is applicable to, e.g., student's t-distribution or to F-distribution of Fisher-Snedecor. As illustrative examples, the paper contains two statistical applications in point estimation of parametric functions. Moreover, it contains a simple proof for the completeness of orthogonal polynomial systems in every Pearson/Ord distribution possessing finite moments of any order.

[28] On matrix variance inequalities
Olkin and Shepp, *J. Statist. Plann. Inference* **130**(2005), 351–358, showed that Chernoff's inequality (for the normal) can be extended to a matrix-valued analogue, and the same is true for the Gamma distribution. In this note we extend these results, obtaining Poincare-type and Bessel-type inequalities for matrices of arbitrary order, and for a large class of continuous and discrete distributions.

[29] Linear estimation of location and scale parameters using partial maxima.
Consider the partial maxima sequence, $X_{n:n} = \max\{X_1, \ldots, X_n\}$, arising from an i.i.d. sequence of r.v.'s $X_1, X_2, \ldots$ with common distribution $F(x; \theta_1, \theta_2) = F_0((x - \theta_1)/\theta_2)$. Here we assume that both parameters, the location parameter $\theta_1 \in \mathbb{R}$, and the scale parameter $\theta_2 > 0$, are unknown. We also assume that F_0 has finite variance. As in the order statistics setup, for estimation purposes we construct the Best Linear Unbiased Estimators, BLUE's; these are estimators of the form $\sum_{i=1}^{n} c_i X_{i:i}$, they are unbiased (for θ_1 or θ_2) and their variance is minimal. However, in contrast to the order statistics setup, the consistency of these estimators is no longer obvious. The reason is that we have a substantial loss of information, since we only record the largest current observation in every step. This work concerns with the point estimation of the scale parameter. The main result provides sufficient conditions in F_0, guarantying that the BLUE of θ_2, $T_2 = T_2^n$, is a weekly consistent estimator of θ_2, that is, $T_2^n \to \theta_2$ in probability, as $n \to \infty$.

We give an idea of the main result in its simplest form: Assume that F_0 has finite second moment and a logconcave density f_0 (or a logconvex density with support bounded from below), and suppose that

$$\lim_{x \to \omega-} \frac{f_0(x)}{(1 - F_0(x))^{\gamma}[-\log(1 - F_0(x))]^{\delta}} = L \in (0, +\infty),$$

where $\omega = \omega(F_0) = \inf\{x : F_0(x) = 1\}$ is the upper end-point of the support of F_0 and γ, δ are constants such that $(\gamma, \delta) \in (-\infty, 3/2) \times \{0\} \cup (1/2, 1] \times (0, +\infty)$. Then, there exists a constant $C = C(F_0)$, such that

$$\mathbb{E}[T_2^n - \theta_2]^2 \leq \frac{C}{\log n}.$$

The conditions are easily verified for several location-scale families that are used in statistics, like Normal, Exponential (Weibull), Logistic, Pareto, Power distribution.

[30] Self-inverse and exchangeable random variables.
An r.v. Z is called self-inverse if the r.v.'s Z and $1/Z$ have the same distribution. It is shown that Z is self-inverse if and only if it can be expressed as $Z = X/Y$ for some exchangeable r.v.'s X, Y.

[31] A simple method for obtaining the maximal correlation coefficient and related characterizations.
The maximal correlation coefficient of X, Y is a classic measure of dependence, defined as

$$R = R(X,Y) = \sup \rho(g_1(X), g_2(Y)),$$

where the supremum is taken over functions $g_1, g_2 : \mathbb{R} \to \mathbb{R}$ such that $0 < \operatorname{Var} g_1(X) < \infty$, $0 < \operatorname{Var} g_2(Y) < \infty$, and $\rho(\cdot,\cdot)$ is the usual Pearson correlation coefficient. It is difficult to calculate R in general, and this is the reason that we know the value of R only in some rare cases; e.g., for the bivariate normal $N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2, \rho)$, $R = |\rho|$.

The present work provides a method to calculate R. The method is applicable when X, Y have finite moments of any order, the polynomials are dense in the corresponding $L^2(X), L^2(Y)$ spaces (e.g, whenever each of X, Y has a finite moment generating function in a neighborhood of zero), and the following *polynomial regression property* holds true:

$$\mathbb{E}(X^n|Y) = A_n Y^n + P_{n-1}(Y), \quad \mathbb{E}(Y^n|X) = B_n X^n + Q_{n-1}(Y), \quad n = 1, 2, \ldots,$$

where P_n, Q_n are polynomials of degree at most n. It is shown that, under the above conditions, $0 \le A_n B_n \le 1$ for all n and

$$R = \sup_{n \ge 1} \sqrt{A_n B_n}.$$

This result provides one-line proofs to known characterizations on order statistics (Terrell, *Ann. Probab.* **11**(1983), 823–826, Székely and Móri, *Statist. Probab. Lett.* **3**(1985), 107–109, López-Blázquez and Castaño-Martínez, *J. Statist. Plann. Inference* **136**(2006), 43–52) and Records (Nevzorov, *Math. Methods Statist.* **1**(1992), 49–54). Moreover, a new characterization of the exponential distribution, based on a novel Splitting-Record model, is given.

[32] Some counterexamples concerning maximal correlation and linear regression.
Due to some old results of Sarmanov (*Dokl. Akad. Nauk SSSR* **121** (1958) 52–55 (in Russian); Mathematical Statistics and Probability, Vol. 2, Amer. Math. Soc., 1962, 207–210; *Dokl. Akad. Nauk SSSR* **120**(1958) 715–718 (in Russian); Mathematical Statistics and Probability, Vol. 4, Amer. Math. Soc., 1963, 271–275; also in Selected Translations), it was believed for some time that whenever X, Y have linear regression to each other, that is, $\mathbb{E}(X|Y) = a_1 Y + a_0$ and $\mathbb{E}(Y|X) = b_1 X + b_0$, then (cf. [31])

$$R(X,Y) = |\rho(X,Y)|.$$

This incorrect implication was used in several papers and books, e.g., Rohatgi and Székely, *J. Stat. Comput. Simul.* **40**(1992), 260–262; Arnold, Balakrishnan and Nagaraja, *Records*, Wiley, 1998, p. 101; Székely and Gupta, *Math. Methods Statist.* **7**(1998), 122;

David and Nagaraja, *Order Statistics*, Wiley, 2003, p. 74; Ahsanullah, *Record Values – Theory and Applications*, Univ. Press Amer. Inc., 2004, p. 23; Barakat, *Arab J. Math.*, **1**(2012), 149–158, to mention a few. In the present note we construct counterexamples to show that the implication is incorrect even if the linear regressions are non-trivial, i.e., $a_1b_1 \neq 0$. One of the examples is just a mixture of two bivariate normals.

[33] An extension of the disc algebra, II
The present work belongs to the area of complex analysis, and concerns uniform approximation of functions by polynomials. We define a new compactification of the complex plane, different of the usual one that uses Riemann's sphere. Under the new compactification, the points at infinity (extended complex numbers, say) have an angle, e.g., $z = \infty e^{i\theta}$, so that the point $z = +\infty$ corresponds to $\theta = 0$, while $z = -\infty$ corresponds to $\theta = \pi$. The distance of two (extended or not) points is defined in terms of the usual distance of their images into the unit disc, according to a suitable homeomorphism:

$$d(z_1, z_2) = \begin{cases} \left|\dfrac{z_1}{1+|z_1|} - \dfrac{z_2}{1+|z_2|}\right|, & \text{if } z_1, z_2 \in \mathbb{C}, \\[2ex] \left|\dfrac{z_1}{1+|z_1|} - e^{i\theta}\right|, & \text{if } z_1 \in \mathbb{C},\ z_2 = \infty e^{i\theta}, \\[2ex] \left|e^{i\theta_1} - e^{i\theta_2}\right|, & \text{if } z_1 = \infty e^{i\theta_1},\ z_2 = \infty e^{i\theta_2}. \end{cases}$$

Let $D = \{z : |z| < 1\}$, $\overline{D} = \{z : |z| \le 1\}$ and $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty e^{i\theta}, \theta \in \mathbb{R}\}$. The main result of the paper shows that a function $f : \overline{D} \to \overline{\mathbb{C}}$ is the uniform limit of polynomials, with respect to the metric d, if and only if

(a) either f is holomorphic in D and continuous (with respect to the metric d) on $\overline{D}$,
(b) or $f(z) = \infty e^{i\theta(z)}$ for some function $\theta : \overline{D} \to \mathbb{R}$ which is continuous (in the usual sense) on $\overline{D}$ and harmonic in D.

[34] Strengthened Chernoff-type variance bounds.
The Pearson family of continuous distributions is defined by

$$\int_{-\infty}^{x} (\mu - t) f(t) dt = f(x) q(x),$$

where $q(x) = \delta x^2 + \beta x + \gamma$ is a polynomial of degree at most 2, f is the density of X and $\mu = \mathbb{E}X$ (assumed finite). This family contains several of the known distributions like Beta, Gamma, Normal, and their negatives. Moreover, if $\delta \le 0$, X has finite moments of any order and the polynomials are dense in $L^2(X)$, see [26], [27]. In the present paper we first define a class of real functions, $\mathcal{H}^n(X)$, as follows: $g \in \mathcal{H}^n(X)$ if

(a) g is defined in the interior of the support of X (which is always an interval),
(b) g is $n-1$ times differentiable and $g^{(n-1)}$ is absolutely continuous, and
(c) $\mathbb{E} q^n(X) g^{(n)}(X)^2 < \infty$.

Next, it is shown that the sequence $\mathcal{H}^n(X)$ is decreasing when $\delta \le 0$. The main result provides an inequality for the variance of $g \in \mathcal{H}^n(X)$ in terms of its derivatives. Specifically, if $\delta \le 0$,

$$\operatorname{Var} g(X) \le \sum_{k=1}^{n} \frac{\mathbb{E}^2 q(X)^k g^{(k)}(X)}{k!\,\mathbb{E} q(X)^k \prod_{j=k-1}^{2k-2}(1-j\delta)} + R_n$$

where

$$R_n = \frac{1}{(n+1)! \prod_{j=n}^{2n-1}(1-j\delta)} \left\{ \mathbb{E} q(X)^n g^{(n)}(X)^2 - \frac{\mathbb{E}^2 q(X)^n g^{(n)}(X)}{\mathbb{E} q(X)^n} \right\}.$$

The sum in the RHS of the inequality (without the remainder R_n) is exactly the lower bound in [27]. Moreover, the inequality can be written in the form

$$\inf \parallel g - p_n \parallel_2 \ \le \sqrt{R_n}$$

where the infimum is taken over polynomials p_n of degree at most n, and $\parallel g \parallel_2 = \big(\int g(x)^2 f(x) dx\big)^{1/2}$. Thus, we have a simple upper bound for the L^2-distance between g and the polynomials, expressed in terms of the derivatives of g. Hence, the result may be useful to the area of numerical analysis.

The main result provides an essential improvement of all previously known results (Poincare-type inequalities) even if $n = 1$. For instance, if X is $N(0, 1)$, the inequality simplifies to

$$\operatorname{Var} g(X) \le \frac{1}{2}\mathbb{E}^2 g'(X) + \frac{1}{2}\mathbb{E} g'(X)^2,$$

while the classic Chernoff inequality (which is valid also for the class $\mathcal{H}^1(X)$) is

$$\operatorname{Var} g(X) \le \mathbb{E} g'(X)^2.$$

The difference of the above bounds is $\frac{1}{2}\operatorname{Var} g'(X) \ge 0$.

[35] Integrated Pearson family and orthogonality of the Rodrigues polynomials: A review including new results and an alternative classification of the Pearson system.
In this work we present an extensive review of the Pearson family of distributions, as they defined in [34]. We describe in detail all distributions that belong to this class, It turns out that there are, essentially, six types of distributions, while the original Pearson's classification contains twelve types. This difference arises from the fact that the integrated family is defined in terms of an integral equation, and not via Pearson's differential equation,

$$\frac{f'(x)}{f(x)} = \frac{p_1(x)}{p_2(x)}, \quad \deg p_i \le i \quad (i = 1, 2).$$

We show these two families are quite different. Specifically, the differential equation may produce some uninteresting densities where, e.g., the Rodrigues polynomials,

$$h_k(x) = \frac{(-1)^k}{f(x)} \frac{d^k}{dx^k}\big(p_2(x)^k f(x)\big),$$

may not be orthogonal with respect to f. One of the results says that the orthogonality of Rodrigues' polynomials is valid if and only if the density f satisfies the integral equation. Other results show that, for a density of the integrated family, useful quantities can be calculated easily, like:

(a) The Fourier coefficients α_n of g with respect to the orthonormal polynomials ϕ_n,

$$\alpha_n = \mathbb{E}\phi_n(X)g(X) = \frac{\mathbb{E}q(X)^n g^{(n)}(X)}{\left(n!\,\mathbb{E}q(X)^n \prod_{j=n-1}^{2n-2}(1-j\delta)\right)^{1/2}}.$$

(b) The leading coefficients

$$\mathrm{lead}(P_n) = \prod_{j=n-1}^{2n-2}(1-j\delta)$$

of Rodrigues' orthogonal polynomials

$$P_n(x) = \frac{(-1)^n}{f(x)}\frac{d^n}{dx^n}\big(q(x)^n f(x)\big).$$

(c) The norm of P_n,

$$\mathbb{E}P_n(X)^2 = n!\,\mathbb{E}q(X)^n \prod_{j=n-1}^{2n-2}(1-j\delta).$$

(d) The quantity $\mathbb{E}q(X)^n$ that appears in (a), (c),

$$\mathbb{E}q(X)^n = \frac{\prod_{j=0}^{n-1}(1-2j\delta)}{\prod_{j=0}^{n-1}(1-(2j+1)\delta)} \prod_{j=0}^{n-1} q\left(\frac{\mu+j\beta}{1-2j\delta}\right).$$

(e) Recurrent relations for the moments,

$$\mathbb{E}X^{n+1} = \frac{(\mu+n\beta)\,\mathbb{E}X^n + n\gamma\,\mathbb{E}X^{n-1}}{1-n\delta}, \quad \mathbb{E}X^0 = 1,\, \mathbb{E}X^1 = \mu,$$

$$\begin{aligned}\mathbb{E}(X-\mu)^{n+1} &= \frac{nq'(\mu)\,\mathbb{E}(X-\mu)^n + nq(\mu)\,\mathbb{E}(X-\mu)^{n-1}}{1-n\delta}, \\ &\qquad \mathbb{E}(X-\mu)^0 = 1,\, \mathbb{E}(X-\mu)^1 = 0.\end{aligned}$$

(f) The k-th derivatives of the orthonormal polynomials are orthogonal polynomials with respect to the density $f_k = q^k f/\mathbb{E}q^k$, and this density belongs to the same type as f. If ϕ_n are the orthonormal polynomials for f and $\phi_{n,k}$ the corresponding for f_k, then the two sets are related through:

$$\frac{d^k}{dx^k}\phi_{n+k}(x) = \left(\frac{(n+k)!\prod_{n+k-1}^{n+2k-2}(1-j\delta)}{n!\,\mathbb{E}q(X)^k}\right)^{1/2}\phi_{n,k}(x).$$

We note that the results (a)–(f) are used in an essential way in [34].

[36] Maximizing the expected range from dependent observations under mean-variance information.
In this work we obtain the optimal upper bound for the expected range,

$$\mathbb{E}R_n = \mathbb{E}(X_{n:n} - X_{1:n}),$$

when $X_{1:n} = \min\{X_i,\ 1 \le i \le n\}$, $X_{n:n} = \max\{X_i,\ 1 \le i \le n\}$, arise from a random vector $(X_1, \ldots, X_n)'$ with known means, $\mu_i = \mathbb{E}X_i$, and known variances, $\sigma_i^2 = \mathrm{Var}\,X_i > 0$. The main result is:

$$\sup \mathbb{E}R_n = \inf_{c\in\mathbb{R},\ \lambda>0} \left\{ -(n-2)\lambda + \frac{\lambda}{2}\sum_{i=1}^{n} U\left(\frac{\mu_i - c}{\lambda}, \frac{\sigma_i}{\lambda}\right) \right\},$$

where the function $U(\cdot,\cdot) : \mathbb{R} \times (0,\infty) \to \mathbb{R}$ is given by

$$U(x,y) = \begin{cases} 2\sqrt{x^2+y^2}, & \text{if } x^2+y^2 \ge 4, \\ 2 + \frac{1}{2}(x^2+y^2), & \text{if } 2|x| < x^2+y^2 < 4, \\ |x| + 1 + \sqrt{(|x|-1)^2+y^2} & \text{if } x^2+y^2 \le 2|x| < 4. \end{cases}$$

For the proof we use convex optimization techniques that extend the proof of Bertsimas, Natarajan and Teo, *Prob. Engineer. Inform. Sci.* **20**(2006), 667–686 and Bertsimas, Doan, Natarajan and Teo, *Math. O. R.* **35**(2010), 580–602, to the case of expected range. Moreover, we characterize all random vectors that attain the equality in the bound (extremal vectors).

The results are based on a key deterministic inequality for the range, ant this inequality can be viewed as the range-analogue of the inequality from Lai–Robbins, *Proc. Nat. Acad. Sci. USA* **73**(1976), 286–288.

Moreover, we provide a detail comparisson of the tight bound with the classic result from Arnold and Groeneveld, *Ann. Statist.* **7**(1979), 220–223, and we characterize the cases where the Arnold–Groeneveld bound is optimal.

[37] On sequences of expected maxima and expected ranges.
Let $X, X_1, \ldots$ be i.i.d. integrable (non-degenerate) r.v.'s and set $X_{n:n} = \max_{1\le i\le n}\{X_i\}$. The problem we study in the present work is the following: Given a real sequence, $\{\mu_n\}_{n=1}^{\infty}$, does there exist an r.v. X such that $\mathbb{E}X_{n:n} = \mu_n$ for all n? Known results relate this question to the well-known *Hausdorff moment problem*, and perhaps, the simplest answer is provided by Kolodynski, *Statist. Probab. Lett.* **47**(2000), 295–300:

The sequence $\{\mu_n\}_{n=1}^{\infty}$ is an expected maxima sequence (EMS for short) if and only if it satisfies the following three conditions.

(a) $(-1)^{k+1}\Delta^k\mu_n > 0$ for all $n \ge 1$ and $k \ge 1$,

(b) $\mu_n = o(n)$ as $n \to \infty$, and

(c) $\sum_{j=1}^{n}(-1)^j\binom{n}{j}\mu_j = o(n)$ as $n \to \infty$.

In practice, however, it is difficult to check the validity of (a)–(c), even for simple sequences like, e.g., $\mu_n = \sqrt{n}$ or $\mu_n = \log(n)$. In the present paper we provide an alternative method, relating the EMS sequences to some Bernstein functions of a particular form, namely we consider those functions $g : [0,\infty) \to [0,\infty)$ that have the form

$$g(x) = \int_{(0,\infty)} (1 - e^{-xy}) d\mu(y), \quad x \geq 0,$$

for some measure μ in $(0,\infty)$ satisfying $\int_{(0,\infty)} \min\{1, y\} d\mu(y) < \infty$.

We prove that the sequence $\{\mu_n\}_{n=1}^{\infty}$ is an EMS if and only if it can be written as

$$\mu_n = \mu_1 + g(n-1), \quad n = 1, 2, \ldots,$$

with g as above and, moreover, the Bernstein function g and the measure μ are unique. Similar results are obtained for sequences of expected ranges. The paper contains simple sufficient conditions and a number of examples that verify the applicability of this representation.

[38] A factorial moment distance and an application to the matching problem.

For two r.v.'s X, Y with values in $\mathbb{N}$ and probability generating functions with radius of convergence $1 + \delta$ (for some $\delta > 0$), we define the *factorial moment distance*,

$$d_\alpha(X,Y) = \sum_{k=1}^{\infty} \frac{\alpha^{k-1}}{k!} \left|\mathbb{E}(X)_k - \mathbb{E}(Y)_k\right|, \quad (\alpha > 0)$$

where $(c)_k = c(c-1)\cdots(k-k+1)$. This distance is useful whenever the factorial moments admit closed forms. Moreover, the inequality

$$d_{TV}(X,Y) \leq d_2(X,Y)$$

is satisfied ($d_{TV}(X,Y) = \sup_A |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|$ the total variation distance), provided that both X, Y have probability generating functions with radius of convergence greater than 2. Some applications to matching problems are given, estimating the rate of convergence to the Poisson distribution.

[39] Orthogonal polynomials in the cumulative Ord family and its application to variance bounds.

The (cumulative) Ord family of distributions (discrete Pearson) contains the integer-valued r.v.'s X with probability mass function p and finite mean μ, satisfying the identity

$$\sum_{j=-\infty}^{k} (\mu - j) p(j) = q(k) p(k), \quad k \in \mathbb{Z},$$

where $q(k) = \delta k^2 + \beta k + \gamma$ is a polynomial of degree at most 2 (see [26], [27]).

In the present work we present a complete classification of these distributions, and we show that there are, essentially, six types. Moreover, we present results regarding the moments and the orthogonal polynomials, similar to those given in [35] for the continuous case. Using these results, we obtain upper and lower bounds for $\mathrm{Var}\, g(X)$, in terms of the forward differences, $\Delta^k g$. The bound for the Poisson distribution with parameter λ reads as follows:

$$(-1)^n \left(\mathrm{Var}\, g(X) - S_{m,n}(g)\right) \geq 0, \quad n = 0, 1, \ldots, \quad m = 0, 1, \ldots,$$

where

$$S_{m,n}(g) = \sum_{i=1}^{m} \frac{\lambda^i}{i!} \frac{\binom{m}{i}}{\binom{m+n}{i}} \mathbb{E}^2 \Delta^i g(X) + \sum_{i=1}^{n} (-1)^{i-1} \frac{\lambda^i}{i!} \frac{\binom{n}{i}}{\binom{m+n}{i}} \mathbb{E}[\Delta^i g(X)]^2.$$

Moreover, the equality in the bound is attained if and only if the function g is a polynomial of degree at most $m+n$. The proposed inequalities provide essential improvements to the results in [26], [27], because they produce strengthened bounds like the ones in [34]. For instance, if $m = n = 1$ and we apply the bound for the Poisson, we get

$$\mathrm{Var}\, g(X) \leq \frac{\lambda}{2} \mathbb{E}^2 \Delta g(X) + \frac{\lambda}{2} \mathbb{E}[\Delta g(X)]^2,$$

while the bound in [26] is

$$\mathrm{Var}\, g(X) \leq \lambda \mathbb{E}[\Delta g(X)]^2.$$

The difference of the above bounds is $\frac{\lambda}{2} \mathrm{Var}\, \Delta g(X) \geq 0$.

(ii) Submitted for Publication

[40] On the limiting distribution of sample central moments.
Let $\{X_i,\ i \geq 1\}$ be i.i.d. non-degenerate r.v.'s with distribution function F and finite $2k$-th moment for some $k \geq 2$. Here we study the asymptotic ($n \to \infty$) distribution of the sample central moments,

$$M_{k,n} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X}_n)^k, \qquad \text{where} \quad \overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

in the particular special case where the limiting distribution is degenerate. Distributions F having this property are called *singular*. It is shown that all singular distributions have two or three supporting points. More specifically, for k even the family of singular distributions contains only two-valued r.v.'s of a particular form, while for odd k, it includes some two-valued and some three-valued distributions. We provide a complete description for all two-valued singular distributions and a partial description for the three-valued ones (which is exact for $k = 3$). Moreover, we prove a characterization of normality which reads as follows: provided that all moments exist, the asymptotic independence of sample mean and every sample central moment characterizes the normal distribution.

Furthermore, using the *Delta* method, we obtain second order asymptotic results for the sample central moments from singular distributions. It is shown that the second order limiting distribution is (a) either a multiple of a χ^2 r.v., or (b) a difference of two multiples of two independent χ^2-r.v.'s with one degree of freedom.

June 2, 2018